Towards Learned Color Representations for Image Splicing Detection

Benjamin Hadwiger*, Daniele Baracchi†, Alessandro Piva†, Christian Riess*

* IT Security Infrastructures Lab,
Friedrich-Alexander University Erlangen-Nürnberg (FAU)

† Dipartimento di Ingegneria dell'Informazione,
Università degli Studi di Firenze (UNIFI)

ICASSP, Brighton, May 2019

Did these events really occur?

Images based on
MS COCO Database

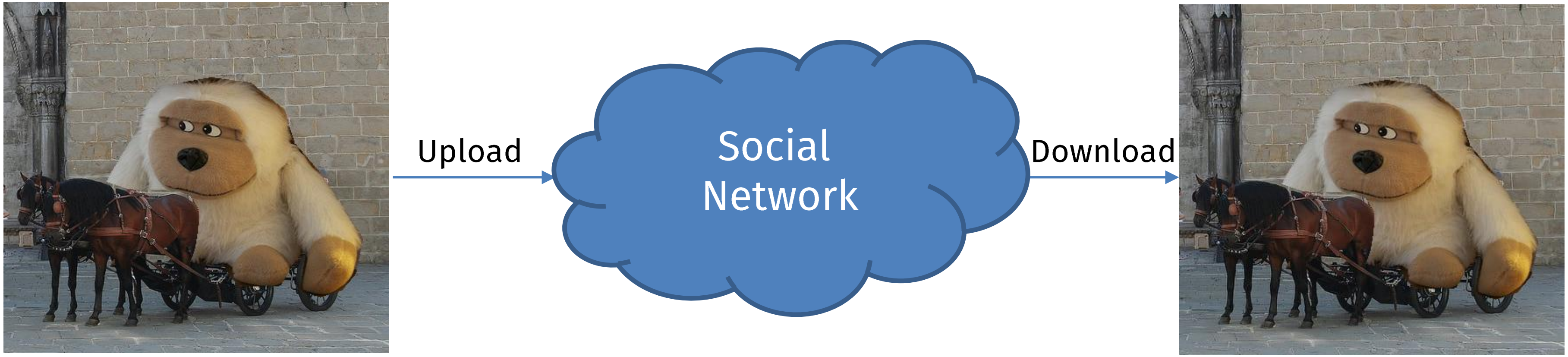
Multimedia Forensics – Goals and Methods

Important goal of multimedia forensics:
Determine authenticity of images

Typical approaches:
Exploit high frequent (HF) image statistics, e.g.

- Camera fingerprint
- Noise statistics
- Compression artifacts
- Resampling artifacts

The Impact of Social Networks



The Impact of Social Networks

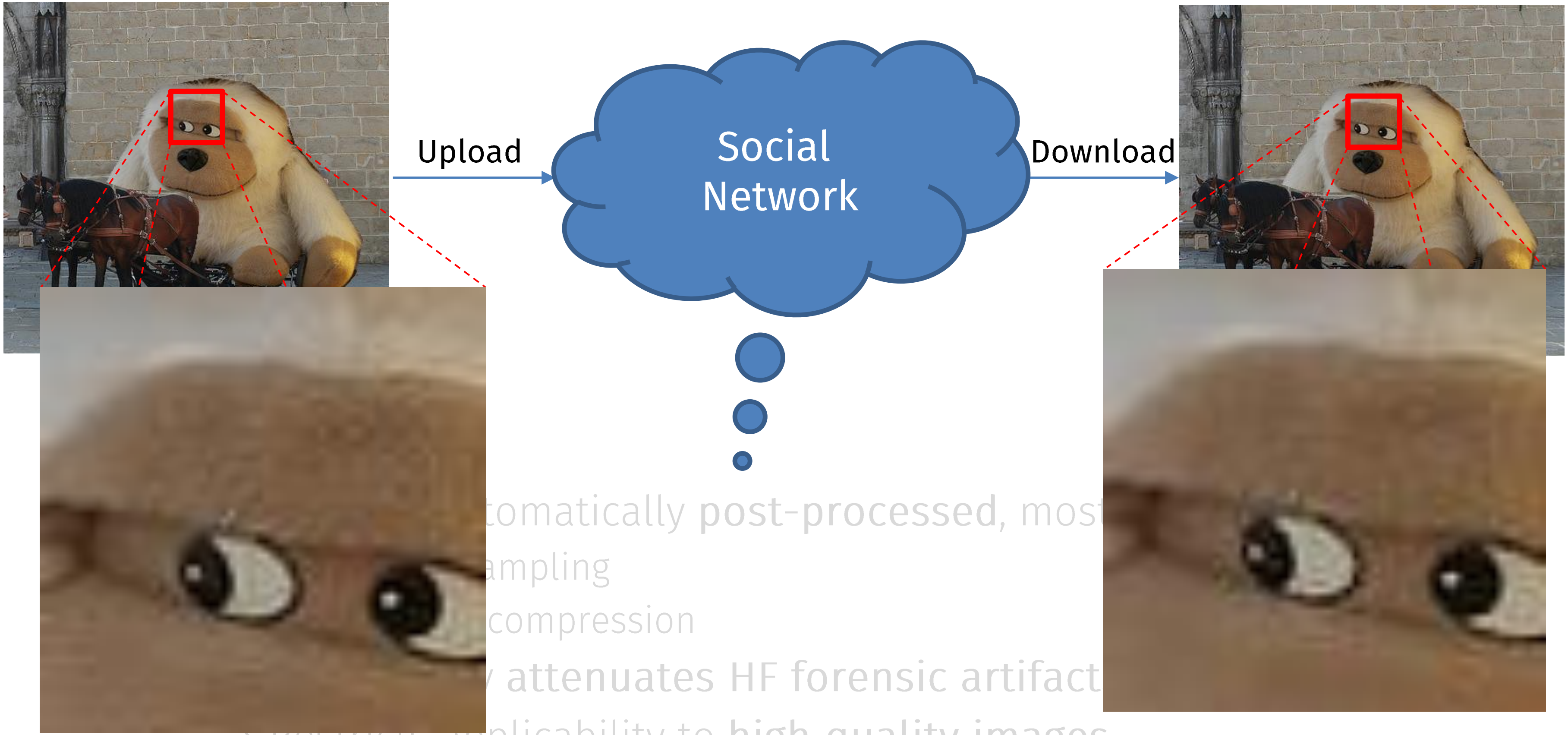
Images are automatically post-processed, most notably:

- Downsampling
- JPEG recompression

→ Significantly attenuates HF forensic artifacts

→ Restricts applicability to high quality images

The Impact of Social Networks



tomatically post-processed, most

ampling

compression

attenuates HF forensic artifact

→ Restricts applicability to high quality images

Towards Robust Manipulation Detection

Can we detect manipulations independently of the image quality?

Towards Robust Manipulation Detection

Can we detect manipulations independently of the image quality?

Physics-based cues are often remarkably robust against post-processing

We explore a novel cue based on the color formation of an image

Towards Robust Manipulation Detection

Can we detect manipulations **independently of the image quality**?

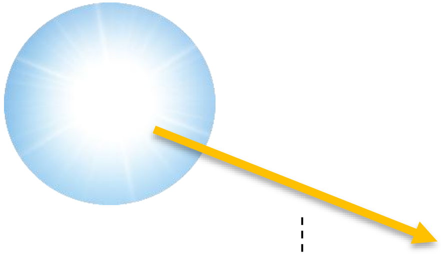
Physics-based cues are often **remarkably robust** against post-processing

We explore a **novel cue** based on the **color formation** of an image

Images based on MIT-Adobe 5k Database

Background: Color Image Formation

Light source

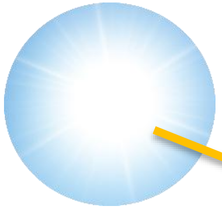


Spectral illuminant density: $e(\lambda)$

$e(\lambda)$

Background: Color Image Formation

Light source



Spectral illuminant density: $e(\lambda)$

$e(\lambda)$

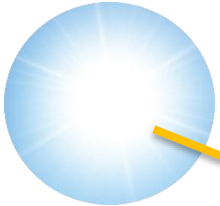
Scene

Spectral reflectance: $r(\lambda)$

$e(\lambda) \cdot r(\lambda)$

Background: Color Image Formation

Light source



Spectral illuminant density: $e(\lambda)$

$e(\lambda)$

Scene

Spectral reflectance: $r(\lambda)$

$e(\lambda) \cdot r(\lambda)$

Camera

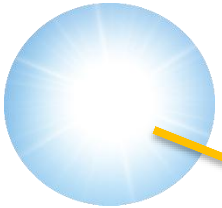
In-camera processing

- Camera sensitivity: $\vec{c}\,(\lambda)$
- White balancing
  Color transformation
  etc.

(White balancing, Color transformation, etc.) } Ω

Background: Color Image Formation

Light source	Scene	Camera	Color image



Spectral illuminant density: $e(\lambda)$

Spectral reflectance: $r(\lambda)$

In-camera processing

- Camera sensitivity: $\vec{c}\,(\lambda)$
- White balancing
  Color transformation
  etc.

(White balancing, color transformation, etc. $\}\ \Omega$)

$e(\lambda)$

$e(\lambda) \cdot r(\lambda)$

$$\Omega\left(\int_{\Lambda} e(\lambda) \cdot r(\lambda) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

Proposed Method: Idea

$$\vec{I} = \boldsymbol{\Omega}\left(\int_{\Lambda} \mathbf{e}(\boldsymbol{\lambda}) \cdot r(\lambda) \cdot \vec{\mathbf{c}}(\boldsymbol{\lambda})\, d\lambda\right)$$

e: illuminant sp. density
Ω: in-camera processing
$\vec{c}$: sp. camera sensitivity
r: spectral reflectance
$\vec{I}$: image intensity

$\mathbf{e}$, $\boldsymbol{\Omega}$ and $\mathbf{c}$ characterize imaging conditions

Assume consistency of $\mathbf{e}$, Ω and $\mathbf{c}$ in pristine image

Image source: NUS Database

Proposed Method: Idea

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

$$\vec{I} = \Omega\left(\int_{\Lambda} e(\lambda) \cdot \mathbf{r}(\lambda) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

e: illuminant sp. density
Ω: in-camera processing
$\vec{c}$: sp. camera sensitivity
r: spectral reflectance
$\vec{I}$: image intensity

e, **Ω** and **c** characterize imaging conditions

Assume consistency of e, Ω and c in pristine image

⬇

How can we control the spectral reflectance $\mathbf{r}(\lambda)$?

Image source: NUS Database

Proposed Method: Idea

$$\vec{I} = \Omega\left(\int_{\Lambda} e(\lambda) \cdot \mathbf{r}(\lambda) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

- e: illuminant sp. density
- Ω: in-camera processing
- $\vec{c}$: sp. camera sensitivity
- r: spectral reflectance
- $\vec{I}$: image intensity

$\mathbf{e}$, $\mathbf{\Omega}$ and $\mathbf{c}$ characterize imaging conditions

Assume consistency of $\mathbf{e}$, $\mathbf{\Omega}$ and $\mathbf{c}$ in pristine image

⬇

How can we control the spectral reflectance $\mathbf{r(\lambda)}$?

⬇

Place Macbeth ColorChecker in image

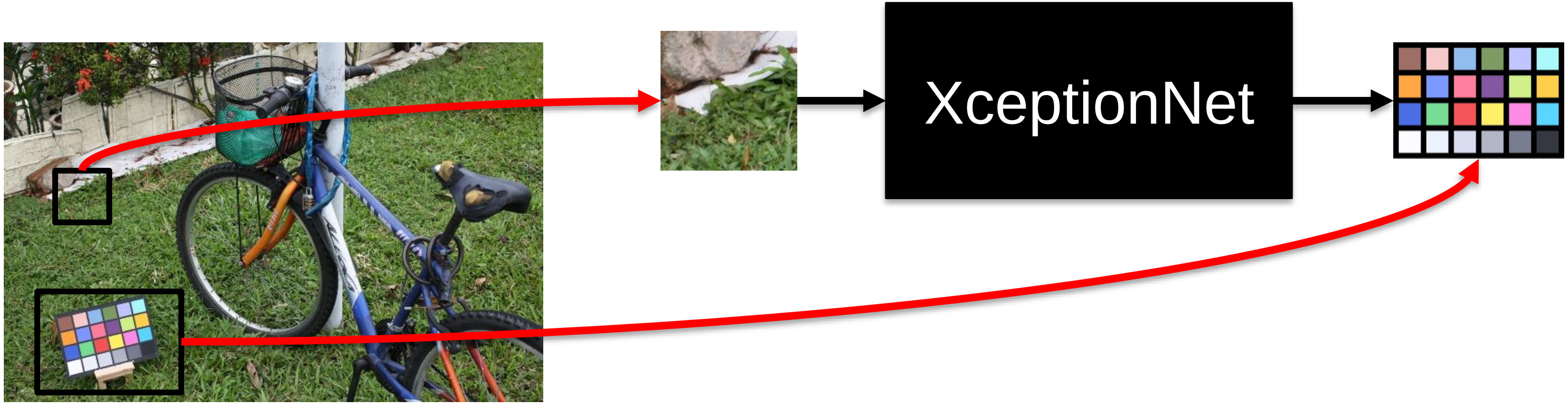
Image source: NUS Database

Proposed Method: Learning the Color Descriptor

1. Train a CNN to locally estimate the observed colors of the ColorChecker

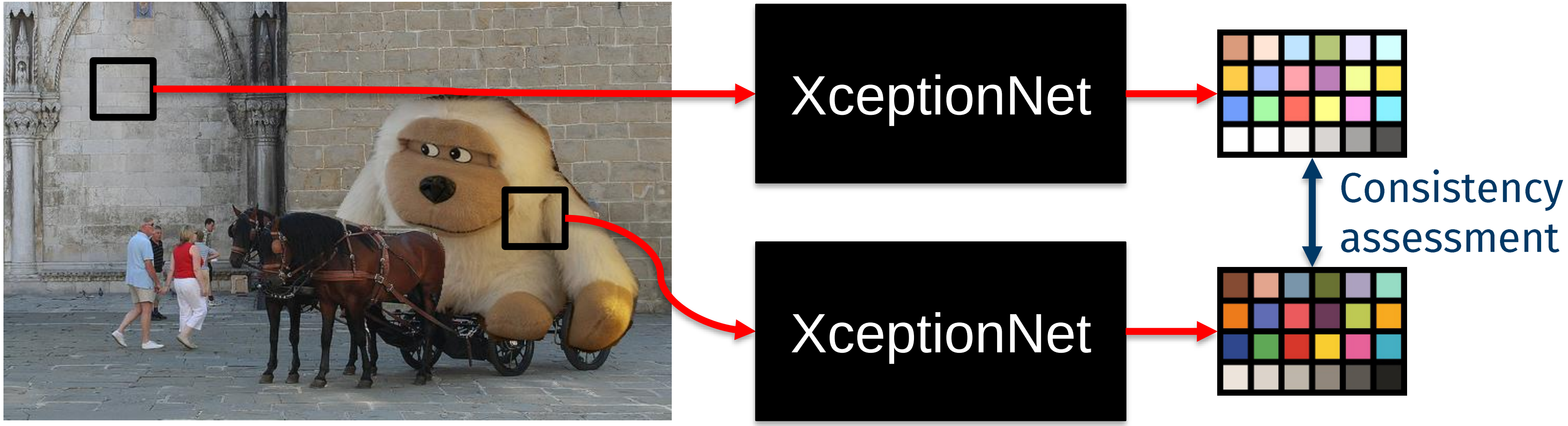
The learned color descriptor is

- Covariant with respect to imaging conditions
- Invariant with respect to reflectance of the image patches



Proposed Method: Consistency Assessment

2. Classify consistency of local estimates



Discriminability of the Learned Color Descriptor

How well do the learned color features characterize the image provenance of a patch?

Discriminability of the Learned Color Descriptor

How well do the learned color features characterize the image provenance of a patch?

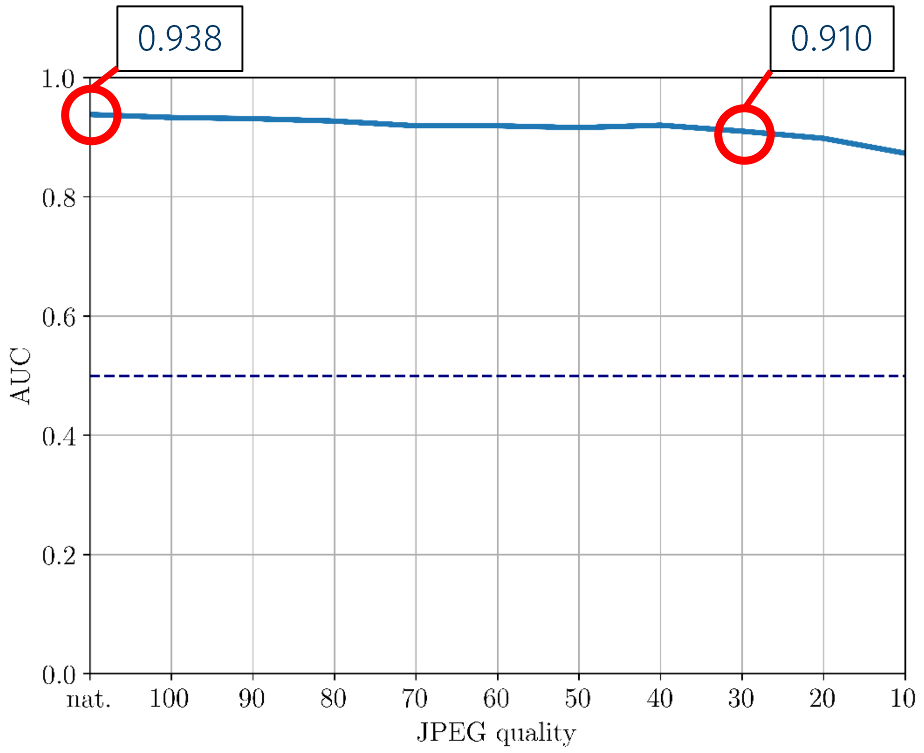
- Extract non-overlapping patches from test images
- Randomly split patches into training / test set
- Train a Random Forest to classify patch provenance: “From which image is this patch?”
- Repeat for increasingly stronger compressions

Discriminability of the Learned Color Descriptor

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

How well do the learned color features characterize the image provenance of a patch?

- Extract non-overlapping patches from test images
- Randomly split patches into training / test set
- Train a Random Forest to classify patch provenance: "From which image is this patch?"
- Repeat for increasingly stronger compressions



Consistency Assessment with the Color Descriptor

A more realistic scenario: Are different regions in an image consistent?

Image based on Dresden Image Database

Consistency Assessment with the Color Descriptor

A more realistic scenario: Are different regions in an image consistent?

- Training on VISION Database,
  Test on Dresden Image Database
- Test splices: combine same scene
  from different cameras
- Assess patch-based consistency
  of reference and target region
- Repeat for increasingly
  stronger compressions

Image based on Dresden Image Database

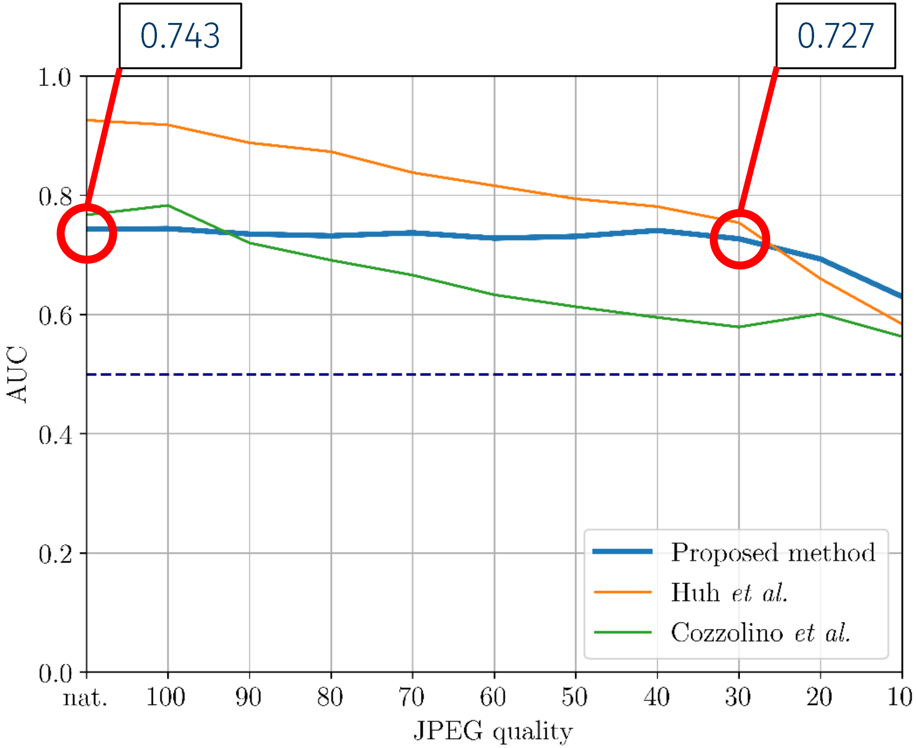
Consistency Assessment with the Color Descriptor

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

A more realistic scenario: Are different regions in an image consistent?

- Training on VISION Database, Test on Dresden Image Database
- Test splices: combine same scene from different cameras
- Assess patch-based consistency of reference and target region
- Repeat for increasingly stronger compressions

Image based on Dresden Image Database



Huh *et al.*: *"Fighting Fake News: Image Splice Detection via Learned Self-Consistency"*, ECCV '18
Cozzolino *et al.*: *"Splicebuster: A new blind image splicing detector"*, WIFS '15

Outlook: Qualitative Results

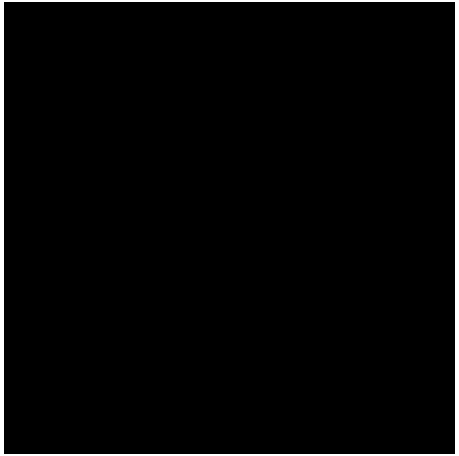
Outlook: Qualitative Results

Outlook: Qualitative Results (*cont.*)

Outlook: Qualitative Results (*cont.*)

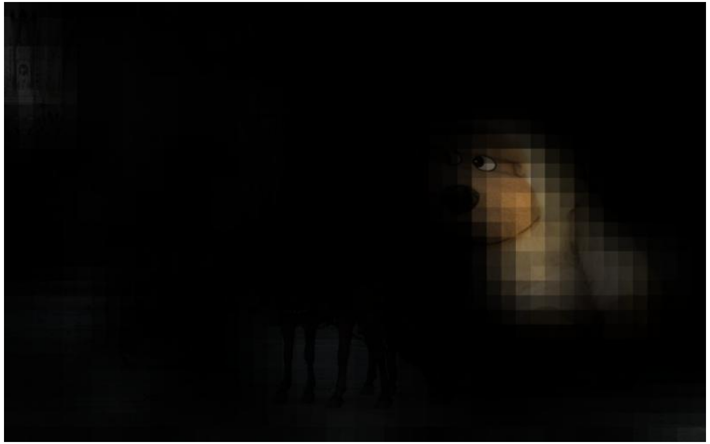
Outlook: Qualitative Results (*cont.*)

Outlook: Qualitative Results (*cont.*)



Outlook: Qualitative Results (*cont.*)

Outlook: Qualitative Results (*cont.*)



Summary

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Conclusion

- We presented a novel cue based on color image formation
- We demonstrated remarkable robustness against JPEG compression
- Promising to work in low-quality settings

Ongoing work

- Incorporate prior knowledge on camera
- Perform consistency assessment using Siamese network

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Thank you!

Questions?